Applications of Quantitative Digital Image Analysis to Breast Cancer Research

C. ORTIZ DE SOLÓRZANO,[1] S. COSTES,[1] D.E. CALLAHAN,[1] B. PARVIN,[2] AND M.H. BARCELLOS-HOFF*
[1]*Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720*
[2]*Computing Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720*

KEY WORDS: mammary gland; carcinogenesis; genomic instability; transforming growth factor-β; ionizing radiation

ABSTRACT Our studies of radiogenic carcinogenesis in mouse and human models of breast cancer are based on the view that cell phenotype, microenvironment composition, communication between cells and within the microenvironment are important factors in the development of breast cancer. This is complicated in the mammary gland by its postnatal development, cyclic evolution via pregnancy and involution, and dynamic remodeling of epithelial-stromal interactions, all of which contribute to breast cancer susceptibility. Microscopy is the tool of choice to examine cells in context. Specific features can be defined using probes, antibodies, immunofluorescence, and image analysis to measure protein distribution, cell composition, and genomic instability in human and mouse models of breast cancer. We discuss the integration of image acquisition, analysis, and annotation to efficiently analyze large amounts of image data. In the future, cell and tissue image-based studies will be facilitated by a bioinformatics strategy that generates multidimensional databases of quantitative information derived from molecular, immunological, and morphological probes at multiple resolutions. This approach will facilitate the construction of an in vivo phenotype database necessary for understanding when, where, and how normal cells become cancer. *Microsc. Res. Tech. 59:119–127, 2002.* Published 2002 Wiley-Liss, Inc.†

INTRODUCTION

The development of cancer requires, first, the subversion and, then, the complicity of normal cells. If we can understand what factors control normal epithelial cells, then we may better understand how cancer develops. Epithelial cells are organized into tissues that dictate cell behavior via the microenvironment, consisting of other cells, insoluble extracellular matrix proteins, soluble hormones, and cytokines (Bissell and Barcellos-Hoff, 1987). Microscopy is the tool of choice to study cells, their relationships and their responses to complex signals from the microenvironment. Currently, image acquisition and analysis is conducted in a linear fashion and information is stored in parameter subsets.

In many tissues of both humans and animals, exposure to high-dose ionizing radiation represents a well-established carcinogen. Our goal is to understand how ionizing radiation leads to breast cancer by studying how mammary epithelial cells respond to the irradiated microenvironment. By using immunofluorescence, digital microscopy, and image analysis to map patterns of radiation-induced protein expression, we have demonstrated that the microenvironment is altered in irradiated tissues, which, in turn, affects the behavior of cells (Barcellos-Hoff, 1998). Tissue response to radiation is rapid, global, tissue-specific, and sensitive. Our current goal is to define the dose response for these features in order to test their dependence on underlying physical and biochemical processes initiated by radiation. To do so, we have developed image analysis strategies to quantify microscopic features that are affected in character, abundance, or intensity after radiation exposure. We envision an integrated approach to imaging information storage and visualization that enables data mining for subsequent analysis of their interrelationships.

CARCINOGENESIS IN EXPERIMENTAL MOUSE MODELS

Epidemiologic data demonstrate that there is an increased risk of breast cancer in women exposed to as little as 1 Gy as a result of atomic bomb (Tokunaga et al., 1991), therapeutic (Mattsson et al., 1993; Mauch, 1995), and diagnostic (Davis et al., 1989) radiation exposures. Radiation has a well-defined physical basis for action and a statistical probability of total and specific chemical events. It is generally thought to produce damage in individual exposed cells at the time of irradiation. However, the behavior of individual cells is dictated by their interactions with each other, such that tissue function is the result of highly coordinated multicellular behavior. When the unit of function is

Abbreviations used: CAM, cell adhesion molecules; ECM, extracellular matrix; FISH, fluorescence in situ hybridization; ROI, region of interest; TGF-β, transforming growth factor-β1; 3D, three-dimensional.

Grant sponsor: Department of Energy, Office of Biological and Environmental Research; Grant Sponsor: DOD-BCRP; Grant numbers: DAMD17-00-1-0306, DAMD17-00-1-0227.

*Correspondence to: M.H. Barcellos-Hoff, Life Sciences Division, Bldg. 74-174, 1 Cyclotron Rd., Lawrence Berkeley National Laboratory, Berkeley, CA 94720. E-mail: MHBarcellos-Hoff@lbl.gov

Received 27 July 2001; accepted in revised form 16 October 2001
DOI 10.1002/jemt.10183
Published online in Wiley InterScience (www.interscience.wiley.com).

multicellular, integration occurs via extracellular signaling through the microenvironment (Bissell and Barcellos-Hoff, 1987). Microenvironment consists of insoluble proteins in the extracellular matrix (ECM), soluble proteins like cytokines, and cell adhesion molecules (CAMs) that link cells to the ECM and each other. The microenvironment is essential for functional organization and differentiation (Barcellos-Hoff et al., 1989; Barcellos-Hoff and Bissell, 1989).

The cell biology of irradiated tissues is indicative of a coordinated damage response program in which individual cell contributions are directed towards repair of the tissue (Barcellos-Hoff, 1998). In normal tissues, a major role of extracellular signaling is inhibition of carcinogenesis by elimination of abnormal cells and suppression of neoplastic behavior. We have proposed that if the surveillance of abnormal cells is disrupted by radiation exposure, then aberrant, potentially genomically unstable, cells can accumulate and proliferate (Barcellos-Hoff and Brooks, 2001).

We tested the contribution of such alterations to the process of carcinogenesis by creating radiation chimeric mammary glands (Barcellos-Hoff and Ravani, 2000). To do so, we took advantage of the postnatal development of the mammary gland and created a gland lacking epithelium. We transplanted unirradiated mammary epithelial cells, which harbored mutant p53 alleles to provide neoplastic potential, to irradiated stroma. Even as long as 14 days after exposure, irradiation dramatically increased the ability of these cells to establish tumors. Thus, when pre-neoplastic mammary epithelial cells are transplanted to an irradiated stroma, the radiation-induced microenvironment promotes the expression of their neoplastic potential (Barcellos-Hoff and Ravani, 2000).

We focused on characterizing radiation-induced remodeling with the goal of future manipulation of critical elements to test their contribution to radiogenic carcinogenesis. An important factor for interpreting the mode of radiation action is the character of the dose response, which requires a quantitative measure of response. In the next sections, we will provide an overview of the biology of irradiated mammary models that we have addressed using microscopy and image analysis.

FEATURE-BASED QUANTITATIVE MICROSCOPY

During a qualitative evaluation of a tissue on a conventional microscope, one can move the plane of focus back and forth in the region of interest to appreciate the character of the staining. This type of qualitative evaluation has been the standard for pathology studies for decades. This human assessment, however, is difficult to reproduce in a quantitative manner for a variety of reasons. First, the expertise of the observers and their knowledge of the histology is a key factor to detect differences between samples. Another important factor is the use of indirect immunofluorescence and its extreme sensitivity to recognize proteins of interest. Finally, most qualitative observations are done from conventional microscopes, which deteriorate the sharpness of the observed optical focal plane by collecting out-of-focus light.

To obtain statistically validated observations, one must move into quantitative evaluation instead of qualitative. The confocal scanning microscope allows the elimination of out-of-focus light by means of a pinhole filter, thus enabling thick specimens to be optically sectioned into a "stack" of 2D images (i.e., slices). The resolution along the optical axis of the confocal microscope is 0.5 μm, which facilitates spatial identification of specific structures by image segmentation. In turn, this permits reconstruction of the major morphological features using 3D visualization software to navigate through the tissue space and get an overall idea of its morphology. As the technology becomes faster and cheaper in conjunction with more sophisticated image analysis, one readily envisions, in the near future, microscopes able to reconstruct the main structures of the tissue in three dimensions in real time to facilitate direct observation.

Quantifying a morphological feature in a tissue is one of the most difficult tasks in image analysis. One needs to identify the specific structure, a step known as segmentation, for accurate quantification. Human eyes are exellent at recognizing spatial features such as structures (objects), shapes, and densities. However, qualitative differences between two samples are often characterized on vague parameters such as "fuzziness," "irregularity," "continuity," "shape," or "texture." It can be difficult to quantitatively measure such qualitative observations. On the other hand, a computer can estimate mean intensity, areas, and volumes very quickly and accurately on large samples, assuming the segmentation was done accurately. Automatic techniques allow a large sampling of a parameter, which, therefore, gives better statistics. Therefore, the ideal analytical tool should take advantage of the incredible pattern recognition ability of the human and the fast and accurate evaluation of the machine to produce statistically testable hypotheses about responses and their relationships within and between samples.

The typical histology of the mammary gland duct and a representative confocal microscope image is shown in Figure 1. One can appreciate the complexity of such a tissue and realize the challenge to quantify its morphological features. One can also understand clearly the need for three-dimensional images to capture the morphology of such tissues. For example, a typical protein structure of interest is the basement membrane, which forms a large network of tubes when visualized by immunostaining in mammary gland sections. This thin layer of proteins separates epithelial cells (or endothelial cells) from the stromal compartment. It is an important mediator of epithelial integrity, primarily serving as a barrier to invasive growth (Lewko et al., 1981; Liotta et al., 1983; Siegal et al., 1981). However, analyzing the membrane on conventional images can be very misleading since the angle of section with respect to the duct adds geometrical artifacts. Three-dimensional analysis with confocal microscopy resolves this issue.

The laminin antibody used in this study produced sharp and clear staining of the epithelial basement membrane. However, it also stained endothelial basement membranes. Therefore, a simple threshold extraction for the segmentation of the basement membrane could not be used. To control for this factor, a

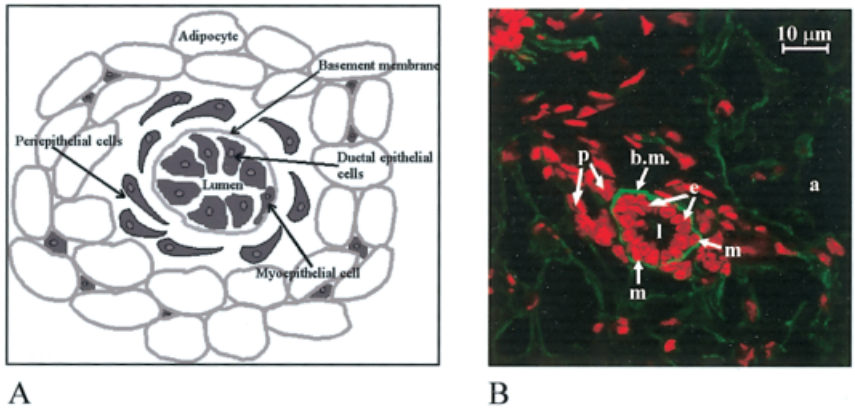

Fig. 1. **A:** The mammary epithelium is organized into a ductal tree that is ensheathed by periepithelial fibrous stroma, embedded in an adipose stroma. **B:** Confocal image of a mouse mammary gland duct using immunofluorescence to detect an ECM protein and a nuclear counterstain (magnification ×800). a: adipocyte; b.m.: basement membrane; e: epithelial cell; m: myoepithelial cell; p: periepithelial fibroblast, 1: lumen [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

smaller region of interest (ROI) was manually defined within the mammary epithelial duct and an automatic isodata threshold technique was applied to create a binary mask of the membrane. This mask was expanded once using a dilation filter, defining a smoother and more continuous membrane, which was used to compute a variety of parameters such as the mean intensity (i.e., average intensity per voxel identified as being in the basement membrane), total number of pixels, contrast, and width.

We used this method to analyze the epithelial basement membrane following densely ionizing radiation. A significant decrease in laminin-1 was both qualitatively and quantitatively evident at 24 hours following radiation exposure (Costes et al., 2000). The data suggest a very specific and fast biological response to particle irradiation, but it is evidently not a gross physical disruption of the basement membrane protein structure. Degradation of laminin following irradiation may alter the ability of epithelial cells to interact with the basement membrane. It is well established that signals generated locally by adhesion receptors themselves are involved in the regulation of cell phenotype (Roskelley et al., 1995). Modification of the laminin network could disrupt the control of epithelial cell growth and differentiation, since signaling via ECM receptors such as integrins intersects with the classic signal transduction pathways. Furthermore, ECM fragments can promote cell motility (Giannelli et al., 1997) and alter gene expression (Stack et al., 1993), possibly by revealing cryptic ECM binding regions (Montgomery et al., 1994), which suggests that a cascade of amplifying events might evolve from localized damage.

ANALYSIS OF EPITHELIAL HETEROGENEITY

Heterogeneity and three-dimensionality are two essential characteristics of any biological system. If understanding heterogeneity is essential for studying normal tissue, it becomes critical in cancer, where tissue heterogeneity and clonal selection are at the very core of carcinogenesis and cancer progression. In breast cancer, little is known about why some parts of the mammary gland are more susceptible than others to develop neoplastic lesions or about what drives the evolution of some neoplastic clones to aggressive phenotypes while their neighbors remain unaltered.

A case in point is the action of transforming growth factor β1 (TGF-β) during mammary gland development and carcinogenesis. We found that TGF-β is a key mediator of the radiation-induced remodeling (Barcellos-Hoff et al., 1994). TGF-β1 is the founding member of a large family of polypeptide growth factors that affect proliferation, programmed cell death (apoptosis), ECM and growth factor production, chemotaxis, and immune response (Roberts et al., 1988). The triple action of TGF-β to inhibit epithelial proliferation, stimulate apoptosis, and mediate ECM deposition and composition suggests that TGF-β is a key regulator of homeostasis and that its inappropriate activation would impact a variety of tissue processes. The biological activity of TGF-β is constrained by its secretion as a latent complex, consisting of TGF-β non-covalently associated with its processed N-terminal pro-segment, called the latency-associated peptide (LAP). Release from LAP is a prerequisite for TGF-β to bind to its cell surface receptors. Events associated with release are called activation, which is considered a critical regulatory event for TGF-β function in vivo (Miyazono and Heldin, 1991; Barcellos-Hoff, 1996). Activation releases TGF-β, which acts as the switch to initiate tissue response to damage in several physiological processes, in particular wounding. Activation has not been widely analyzed temporally or spatially in vivo. TGF-β action has been implicated in certain events in vivo indirectly using experimental manipulations such as neutralizing antibodies or transgenic manipulation leading to overexpression or knockout.

We developed an immunostaining protocol to preserve endogenous latent TGF-β and to discriminate between latent and active TGF-β by immunohistochemistry (Barcellos-Hoff, 1996). We found that latent TGF-β is abundant in normal mammary gland but active TGF-β is restricted to epithelium. This pattern of TGF-β immunoreactivity alters significantly 1 hour after radiation exposure (Barcellos-Hoff et al., 1994). Active TGF-β immunoreactivity was dramatically increased in the epithelium and stroma concomitant with decreased LAP immunoreactivity, which is consistent with a process in which LAP degradation reveals previously masked regions of TGF-β. This pattern persists for more than 7 days after radiation, suggesting that there is a chronic stimulus for TGF-β activation.

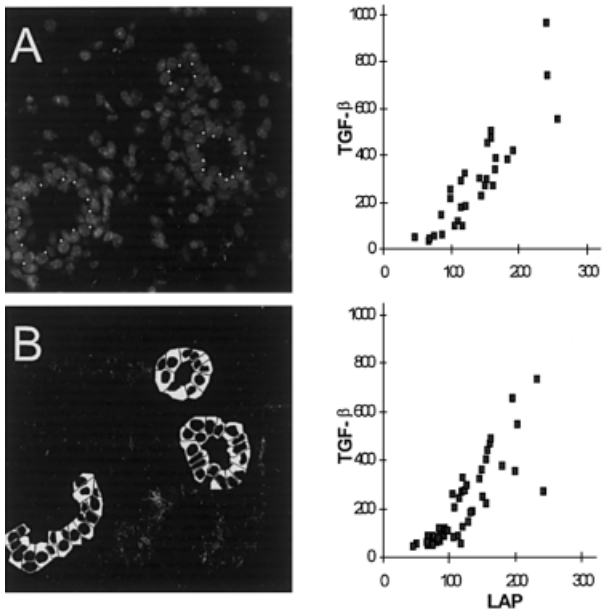

Fig. 2. Methods for measuring relative fluorescence in representative ROI (**A,** white dots on DAPI image) or manual segmentation (**B,** mask) within the cytoplasm of mammary epithelial cells. Data are plotted for each method in graphs on the right of the relative intensity of LAP vs. TGF-β immunoreactivity. Similar data are obtained from ROI as from manual segmentation, but with less labor.

To further understand TGF-β's role in mammary gland biology, it is critical to understand when and where activation occurs during mammary gland development (Barcellos-Hoff and Ewan, 2000). We have recently found that TGF-β activation occurs in a subset of cells during periods of proliferation, which is surprising given its role as a growth inhibitor (Barcellos-Hoff and Ewan, 2000). We wanted to ascertain the intensity of staining on a per cell basis without the laborious, and potentially biased, manual segmenting of individual cells. Therefore, an observer positioned a small region of interest between epithelial nuclei using the nuclear-stained image, without reference to the immunostained images to avoid selection bias. Total fluorescence intensities for each image layer were then generated. When this approach was compared to manual segmenting, the data were virtually identical (Fig. 2).

Further investigation has revealed that this subset consists primarily of steroid receptor (i.e., estrogen and progesterone) positive cells, which are heterogeneously distributed in the mammary epithelium as a function of development and hormone exposure (Shyamala et al., 1997). The heterogeneity of the mammary epithelium is a major factor in understanding the biology of the mammary gland and breast cancer. The relationship between genetic and phenotypic characteristics of intact cells and their local tissue environment is likely to provide insight in the mechanism of cancer initiation and progression. However, many standard analysis methods in molecular biology cannot account for intercellular variation, in that they provide average information from bulk cell populations, taken outside their native tissue context. Analytical methods based on immunohistochemistry and two-dimensional microscopy can account for heterogeneity but neglect or seriously compromise three-dimensionality by extrapolating from a flat projection of a three-dimensional reality. Thus, large-scale morphological patterns, such as distance from primary ducts, or relationships between neoplastic foci, are lost.

Tools to integrate three-dimensional information from tissue at multiple levels of resolution would be of benefit for understanding how cells behave as a function of changing microenvironments within the same tissue. We have developed quantitative analysis of genetic and phenotypic characteristics based on three-dimensional imaging at the cellular level. Local, high-resolution analysis within small tissue volumes of thick tissue sections analysis begins with segmenting (i.e., delineating) every cell in images obtained by confocal microscopy. Cell nuclei stained with a fluorescent DNA-binding counterstain produce high-contrast images. Previous methods use interactive manual or automatic algorithms that provide high performance but are slow and laborious. They are only practical in situations where only few nuclei per sample, or a few samples, must be analyzed. Fully automated algorithms, on the other hand, are much faster, enabling the analysis of hundreds of nuclei. However, although their performance is high (>90%) for specimens containing isolated nuclei, it significantly deteriorates for highly clustered samples, e.g., solid tumor specimens, where the cells are structurally dominated by their nuclei, leaving little separating cytoplasm.

Our approach combines the speed of automation with the accuracy of the human visual system for identifying and classifying objects (Ortiz de Solórzano et al., 1999). Images are pre-segmented by using an automatic adaptive threshold that extracts the stained nuclear volumes from the background. Then, the nuclear volumes are rendered in 3D and presented to the user one by one for interactive classification. The user counts with a set of interactive tools that simplifies and speeds up the process (Fig. 3). Finally, clusters are automatically divided into parts, which are visually classified again and inappropriately divided parts rejoined. Although the interactive classification increases the amount of human effort in an otherwise automatic procedure, it ensures that only nuclei that were correctly segmented based on visual judgment are used for obtaining results. The output is a three-dimensional virtual reconstruction of the tissue, where each nucleus is precisely located and its morphological parameters are known (Fig. 4). This algorithm provides very high segmentation rates, similar to the manual algorithms, while keeping the interaction demand to a minimum.

Intranuclear elements such as copies of genes, RNA, or nuclear proteins can be labeled using fluorescence in situ hybridization (FISH) or immunohistochemistry respectively. Quantification of these elements after nuclear segmentation can be done using multicolor image acquisition and analysis. The analysis may consist of automatically segmenting the copies of the genes targeted by FISH (Fig. 3) or quantifying the amounts of the protein. Among other applications, we have used our system to quantify the level of genetic instability and its progression in human breast cancers (Ortiz de Solórzano et al., 1999b). Genetic instability, which is a

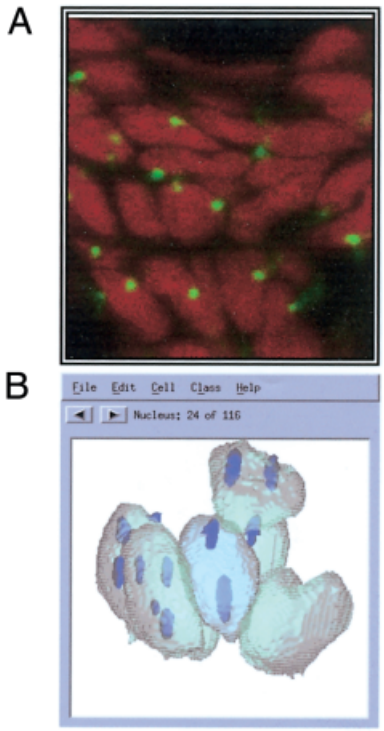

Fig. 3. Segmentation of FISH spots. **A:** 2D section from a 3D confocal image taken on a block of normal mammary gland. FISH was performed with a probe targeting the centromere of chromosome 1 (bright spots) and the nuclei were counterstained with a DNA dye. **B:** The DaVinci program incorporates the segmented FISH signals to the 3D virtual reconstruction of the tissue based on the nuclear segmentation. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

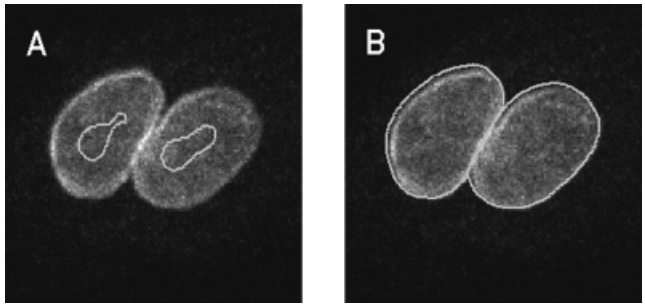

Fig. 4. Segmentation of nuclei using nuclear membrane related proteins. **A:** 2D confocal image from a culture of epithelial cells stained with an antibody against a component (lamin A) of the nuclear lamina. The overlaid red lines show the seeds automatically found by our segmentation algorithm that identify each on the nuclei. **B:** Segmentation obtained after applying the algorithm referenced in the text.

hallmark of breast cancer, is believed to cause the genetic "diversity" required for the accumulation of genetic abnormalities that drives cancer progression. The very concept of genetic instability includes inter- and intracellular variability of chromosome copy number that can only be quantified using per cell analysis. We measured chromosome copy number variability in breast tumors at various histologically defined areas within normal, preneoplastic, and neoplastic breast lesions. Using this program, we found an unexpectedly high level of genomic instability in samples of ductal carcinoma in situ and invasive carcinomas. In addition, we noted variation in the pattern of instability between neighboring areas that could have been missed using other methods.

Because the cells are so highly packed in advanced cancer specimens, there is no or very little background between nuclei to allow our segmentation algorithm to correctly delineate the individual nuclei (Ortiz de Solórzano et al., 2001). In such cases, we used an alternate avenue by staining proteins located on the nuclear lamina, a protein mesh that underlies the inner membrane of the nuclear envelope. By using fluorochrome-tagged antibodies against one or a combination of lamins, a detectable signal between nearest neighbor nuclei can be obtained. Thus, the domain occupied by each nucleus is nicely delimited, no matter how clustered the nuclei might be. The segmentation method requires finding an internal seed that uniquely identifies each cell/nucleus in the image. Then, the surface of each seed is dilated until it aligns with the nuclear lamina. The movement of the surface is geometrically driven by the characteristics of the image (Fig. 4).

Because the automatic delineation and classification of nuclei is an important step in mapping cell responses, in the absence of a nuclear lamin staining, we have developed another approach from grouping roof edges (boundaries between nuclei) and step edges (boundaries between nuclei and background) to a novel solution that we call "regularized centroid transform" (Cong and Parvin, 2000). The model assumes that nuclei are convex or they can be parameterized into a quadratic form (at least locally), and the grouping involves combining partial computed features into a globally consistent representation (Parvin et al., 2000).

However, for studying tissue variations within large volumes or tissue, such as the whole mammary gland of a mouse or a complete tumor biopsy, a confocal approach produces immense amounts of data that are difficult to handle. Thus, we have developed a quantitative system that allows low-resolution virtual 3D reconstruction of the tissue from serial sections, followed by high-resolution cell-level reconstruction and quantitative analysis of the structures of interest. This distributed microscopy imaging system allows acquisition and registration of low magnification (1 pixel = 5 μm) conventional, bright field, or fluorescence images of entire tissue sections. We have used it to trace the ducts of the mammary gland from adjacent sections to create a 3D virtual reconstruction of the gland. Then, the system facilitates revisiting areas of interest for high-resolution (1 pixel = 0.5 μm) imaging and automatic analysis. The system is based on a client-server architecture that permits users to request microscope functions. Some are basic microscope operations (setting the objective lens, exposure time, fluorescence filter; focusing; acquiring single images), while other are more complex, such as scanning multiple field-of-view areas or selecting areas of interest from previously acquired images. Several basic data analysis or management operations are provided, such as creating

cases (i.e., related sets of images), adding whole sections or areas of interest to a case, registering consecutive sections, marking and linking structures of interest on consecutive sections.

The first step of 3D reconstruction of a small volume of tissue is the low-resolution acquisition of images of the entire section for a sequential series. After registering these images, specific features, such as the position of ducts, are marked using a wide variety of indicators such as points, lines, or even shapes. Then, the gross tissue morphology is reconstructed from the markings. High-resolution analysis involves automated revisiting, acquiring, and analyzing areas of interest defined by the user on the 3D reconstruction. Examples of analysis are segmentation of nuclei, detection of intranuclear elements such as genes using FISH, or distribution of immunolabeled proteins.

The strengths of this system are mainly based upon the use of Java at the client side. This provides a universal tool that can be used from any web browser, on any platform, anywhere. The use of separate threads for microscope-linked and non-microscope-linked operations provides a flexible and robust application that allows the sharing of images and resources from different research facilities given its distributed character. The system is now being applied to study the distribution of hormone receptors throughout the ductal tree of the mammary gland to understand the underlying topological organization that is difficult to appreciate when viewed in 2 dimensions. In future studies, we will use this tool to define the cellular composition of gross anatomical features such as ductal branching, which in turn will provide insight into subtle alterations in branching patterns in transgenic mammary models and during neoplastic progression.

MODELING CARCINOGENESIS IN HUMAN CELLS

We have recently begun studies of radiation-induced microenvironment interactions in cultured human mammary epithelial cells. Classical cancer research has benefited from monolayer culture assays for immortality, transformation, and progression (Harris, 1987; Nettesheim and Barrett, 1985), but when any cell type is isolated and established in standard monolayer culture, many of the cell- and tissue-specific traits are lost (reviewed in Bissell, 1981). Indeed, it is difficult to even distinguish by morphology or behavior the differences between normal and malignant human breast cells in a tissue culture dish (Petersen et al., 1992).

Functional culture systems to study regulation of tissue-specific gene expression have been established using cells cultured within a 3-dimensional ECM (Barcellos-Hoff and Bissell, 1989; Bissell and Barcellos-Hoff, 1987; Roskelley et al., 1995). This experimental model demonstrated that differentiated function and morphogenesis typical of the tissue of origin requires interactions between cells and ECM. Thus, mouse and human breast epithelial cells organize into glandular acini, consisting of polarized cells forming a discrete lumen and functionally differentiated surfaces. Tissue-specific organization and polarity are maintained by appropriately localized adhesion molecules, including intercellular E-cadherin, basal-lateral β1-integrin, and basal α6-integrin (Weaver et al., 1997).

Furthermore, three-dimensional morphogenesis within reconstituted basement membrane readily distinguishes between the behaviors of tumorigenic and non-tumorigenic mammary epithelial cells, which are nearly indistinguishable when cultured as monolayers (Petersen et al., 1992). Tumor cells continue to proliferate and fail to establish appropriate cell-cell and cell-ECM connections but non-malignant mammary epithelial cells have growth arrest and form acini similar to those found in situ. The importance of appropriated cellular interactions for modeling radiation effects is underscored by observations that the same cells cultured in 2 (i.e., monolayers) vs. 3 dimensions (e.g., multicellular spheroids) differ in their responses to ionizing radiation, viral infection, cytotoxic drugs, and chemotherapeutic agents.

We evaluated the response to ionizing radiation in terms of this ability to organize into acini as a functional endpoint of cell-cell and cell-ECM interactions. Preliminary studies reveal that irradiated pre-malignant human mammary cells fail to establish polarity, display pronounced acinar disorganization, and exhibited defective expression of cell adhesion molecules (CAM) and gap junction proteins necessary for tissue-specific organization. Unlike our studies in the quiescent adult mammary gland, the establishment of acinar colonies in the human cell culture model requires proliferation. Thus, this study indicates that irradiated cells give rise to progeny that maintain an altered phenotype long after radiation exposure.

The significance of the irradiated phenotype is suggested by a variety of studies showing that loss of microenvironment constraints has profound consequences on tumorigenesis, progression, and metastasis (Bissell et al., 1999; Tlsty, 1998). Our studies show that the disruption of extracellular interactions occurs globally in mouse tissue and human cell populations following irradiation, which clearly limits the role of mutational mechanisms. Phenotypic evolution is a more likely basis of the irradiated phenotype. Phenotype is driven by biochemical changes, due in part to extracellular signaling and epigenetic modulation of the genome. This can lead to heritable phenotypes (as evidenced by the differentiation of more than 300 cell types from a human genome) or a reversible phenotype, such as in the "activated" phenotypes that are observed in certain cells during inflammation or wound healing.

Indeed, the challenge of the post-genomic era is functional genomics, i.e., understanding *how* the genome is expressed to produce myriad cell phenotypes. A phenotype is the result of selective expression of the genome in response to the microenvironment. To use genomic information to understand the biology of complex organisms, the biological responses and signaling pathways in cells need to be studied in context, i.e., within a proper tissue structure. In turn, this information can then be used to understand the development of disease and to more accurately predict risks associated with exposure to radiation or other insults. Quantitative analysis of phenotype requires the microscopic evaluation of many samples so that responses may be classified by degree and type as a function of dose and time after exposure in various models. Microscopy allows observation of specific proteins using antibodies and other reagents, their distribution and abundance, and the spatial determinants, whether subcellular, cellu-

lar, or multicellular, within a sample. We will discuss new approaches for quantitative analysis of interrelated image features to create a more detailed description of phenotype that, in turn, can lead to hypothesis generation.

IMAGING BIOINFORMATICS

To the non-expert, images, although provocative, lack features that are readily interpreted as critical information. The key to a microscopist's skill is that context imparts a great deal of information because images don't exist in a vacuum; images are about something. Thus, image analysis is most useful when it is based on the ability of an informed observer to identify specific features for measurement, that in turn, makes the information interpretable to the uninitiated and allows statistical testing of its importance.

Another significant aspect of microscopy of fixed specimens is that changes in shape, response, and organization in a population of cells, whether in vivo or in culture, are distributed statistically and that observations cannot take place on the same sample over time. It is, therefore, necessary to conduct large population studies and correlate distinct features measured from images with annotation data in order to fully characterize a pattern of response. Thus image analysis resulting in segmentation of spots and dots that correspond to nuclei and chromatin is further enhanced by features such as size, frequency, location relative to other features, attributes, and relation to other images in a data set.

Quantitation of spatial and temporal patterns of multiple markers in both tissues and 3-dimensional cell culture is hampered by labor-intensive methods, a lack of quantitative tools, and the inability to index information. Ideally, one would track fluorescence probes for multiple target proteins, their cellular context, and morphological features in large samples. Currently, sequential measurements obtained with different microscopy techniques preclude detailed analysis of multidimensional responses (e.g., time and space). This type of cataloguing also fails to relate features in novel and meaningful ways that will further our understanding of the basic biology of breast cancer.

To address these issues, we have developed a bioinformatics approach to image acquisition, analysis, and hierarchical image abstraction to create a database that registers localization and expression information about multiple targets along with positional references and morphological features (Fig. 5). A program, called BioSig, stores data in an integrated format to allow the development of tools to query the data for common patterns, relationships between different tissues probed with the same antibodies, and methods to relate information from samples imaged with conventional, deconvolution, and confocal microscopes (Parvin et al., 2000). Image analysis tools for complex pattern recognition can (1) eliminate and reduce manual measurements over large data sets; (2) improve quality control through reducing inter-observer as well as intra-observer variation; (3) provide robust measurements for statistical tests of hypotheses; and (4) allow data mining and visualization on a multidimensional database.

The informatic system consists of three components: (1) data model, (2) presentation manager, and (3) query manager. These subsystems are decoupled for ease of development, testing, and maintenance. The data model captures required annotation data and couples them to a computed representation of images so that hypotheses about the relationships between features can be tested. The model is object oriented and allows bidirectional tracking of annotation and measured feature data. The presentation manager provides mapping between the data model and the user interface. The query manager maps high-level user queries to the Java objects that implement the data model. Thus, the detailed manipulations of the database are hidden, which simplifies the application of analytic tools for the end user.

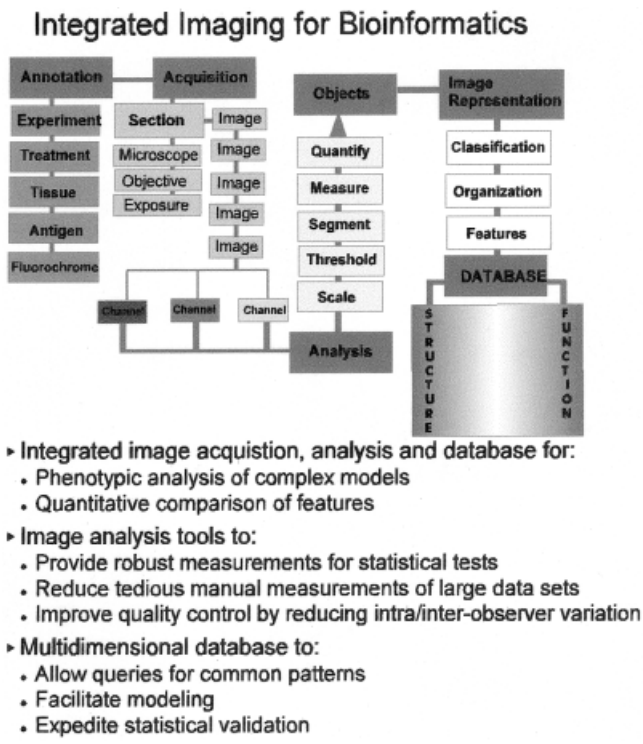

Fig. 5. Schematic of bioinformatics framework for microscopy that integrates image acquisition, annotation, analysis, and representation.

A system named DeepView, developed for accessing and operating microscopes in a collaborative fashion (Parvin et al., 1999), is a "microscopy channel" over the wide area network. The channel is a collaborative problem-solving environment that allows for both synchronous and asynchronous collaboration. Users can seamlessly participate in a microscopy experiment, acquire expert opinions, collect and process data, and store this information in an electronic notebook. The current testbed includes several unique electron and optical microscopes with applications ranging from material science to cell biology.

DeepView consists of four basic services built on top of defined services. Instrument services provides for any type of microscope control. Exchange services provides a common set of utilities for management, shared instrument access, and generic graphical user interface

operation. Declarative services provides flow control for improving the dynamic behavior of the collaborative infrastructure. Computational services provides the analytical capabilities needed for online microscopy and a problem-solving environment. However, DeepView provides a "flat file" system of logging and storing annotation and image data, which is not adequate for the efficient access of large-scale data in a systematic way. A new addition to DeepView is an informatic framework, which includes an object-oriented database for storage and retrieval (Parvin et al., 1999). The database interface is via a web browser; however, all image analysis operations are performed with the computation services.

Data models link image analysis to detailed experimental design and protocol information. We used a data model that is object-oriented and links a particular project to computed features from a collection of images. This link is bidirectional to allow tracking of information from any end point. Each project has its own database, which is linked to studies, consisting of in vivo or in vitro experiments in which samples are imaged and pertinent features are computed. A presentation manager is used to browse the database and visualize the result of a query function. Browsing the database is performed against a predefined schema that captures annotation data, images, and corresponding features. As discussed above, an informed user designates specific features for quantitation as the basis for selecting and development of image segmentation tools. Together, these features are the basis for counting cells, classifying nuclei, measuring protein expression, and quantifying organization of multicellular systems. The resulting data constitute a description of the range of phenotypes generated from a given genome in an organism or model.

We have developed BioSig for studying phenomics, i.e., the biological characteristics resulting from the expression of a specific genome. Imaging bioinformatic databases that register localization information about multiple targets (e.g., growth factors, cell cycle determinants, cell adhesion, extracellular matrix) along with positional references and morphological features permit user interrogation to test new hypotheses regarding functional relationships between cellular responses to radiation. Used in conjunction with animal models and human cell culture models, this novel approach will be used to construct an "irradiated phenotype" database, that in turn will be used to identify critical biological responses to low-dose radiation exposure that can be employed in computational models of radiation risk. Its application to breast cancer histology and pathology in a large population could also be used to generate a "breast cancer" phenotype database to reinforce and quantify the link between genomics (genetic background and mutation), phenomics (pathology), and biological response (prognosis and therapeutic response).

ACKNOWLEDGMENTS

Support for this work was provided by a grant to M.H.B.-H. and B.P. from the Department of Energy, Office of Biological and Environmental Research under contract to the Lawrence Berkeley National Laboratory. C.O.D.S. acknowledges the support of DOD-BCRP projects DAMD17-00-1-0306 and DAMD17-00-1-0227.

REFERENCES

Barcellos-Hoff MH. 1996. Latency and activation in the regulation of TGF-β. J Mammary Gland Biol Neoplas 3:353–363.

Barcellos-Hoff MH. 1998. How do tissues respond to damage at the cellular level? The role of cytokines in irradiated tissues. Radiation Res 150:S109–S120.

Barcellos-Hoff MH, Bissell MJ. 1989. Mammary epithelial cells as a model for studies of the regulation of gene expression. Functional epithelial cells in culture. New York: Alan R. Liss, Inc. p 399–433.

Barcellos-Hoff MH, Brooks AL. 2001. Extracellular signaling via the microenvironment: A hypothesis relating carcinogenesis, bystander effects and genomic instability. Radiat Res 156:618–627.

Barcellos-Hoff MH, Ewan KB. 2000. TGF-β and mammary gland development. Breast Cancer Res 2:92–100.

Barcellos-Hoff MH, Ravani SA. 2000. Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells. Cancer Res 60:1254–1260.

Barcellos-Hoff MH, Aggeler J, Ram TG, Bissell MJ. 1989. Functional differentiation and alveolar morphogenesis of primary mammary epithelial cells cultures on reconstituted basement membrane. Development 105:223–235.

Barcellos-Hoff MH, Derynck R, Tsang ML-S, Weatherbee JA. 1994. Transforming growth factor-β activation in irradiated murine mammary gland. J Clin Invest 93:892–899.

Bissell MJ. 1981. The differentiated state of normal and malignant cells or how to define a "normal" cell in culture. Int Rev Cytol 70:27–100.

Bissell MJ, Barcellos-Hoff MH. 1987. The influence of extracellular matrix on gene expression: Is structure the message? J Cell Sci 8:327–343.

Bissell MJ, Weaver VM, Lelièvre SA, Wang F, Petersen OW, Schmeichel KL. 1999. Tissue structure, nuclear organization, and gene expression in normal and malignant breast. Cancer Res 59 (Suppl 7):1757–1763s; discussion 1763s–1764s.

Ortiz de Solórzano C, Chin K, Knowles D, Jones A, Garcia E, Gray JW, Lockett SJ. 1999. 3D confocal microscopy and image analysis for measurement of genetic instability. Microscopy and Microanalysis 5(s2):1022–1023.

Cong G, Parvin B (2000). A new regularized approach for contour morphing. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA. 1:458–463.

Costes SV, Streuli CH, Barcellos-Hoff MH. 2000. Quantitative image analysis of laminin immunoreactivity in 1 GeV/amu iron particle irradiated skin basement membrane. Radiat Res 154:389–397.

Davis FG, Boice JD, Hrubec Z, Monson RR. 1989. Cancer mortality in a radiation-exposed cohort of massachusetts tuberculosis patients. Can Res 49:6130–6136.

Giannelli G, Falk-Marzillier J, Schiraldi O, Stetler-Stevenson WG, Quaranta V. 1997. Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. Science 277:225–228.

Harris CC. 1987. Human tissues and cells in carcinogenesis research. Cancer Res 47:1–10.

Lewko W, Liotta LA, Wicha MS, Vonderhaar BK, Kidwell WR. 1981. Sensitivity of N-nitrosomethylurea-induced rat mammary tumors to cis-hydroxyproline, an inhibitor of collagen production. Cancer Res 41:2855–2862.

Liotta LA, Rao CN, Barsky SH. 1983. Tumor invasion and the extracellular matrix. Lab Invest 49:636–649.

Mattsson A, Ruden B-I, Wilking N, Rutqvist LE. 1993. Radiation-induced breast cancer: long-term follow-up of radiation therapy for benign breast disease. J Natl Cancer Inst 85:1679–1685.

Mauch P. 1995. Second malignancies after curative radiation therapy for good prognosis cancers. Int J Radiat Oncol Biol Phys 33:959–960.

Miyazono K, Heldin C-H (1991). Latent forms of TGF-β: molecular structure and mechanisms of activation. In: Bock GR, Marsh J, editors. Clinical applications of TGF-β. Chichester: John Wiley. 157:81–92.

Montgomery AM, Reisfeld RA, Cheresh DA. 1994. Integrin alpha v beta 3 rescues melanoma cells from apoptosis in three-dimensional dermal collagen. Proc Natl Acad Sci USA 91:8856–8860.

Nettesheim P, Barrett JC. 1985. In vitro transformation of rat tracheal epithelial cells: a model for the study of multistage carcinogenesis. Carcinog Compr Surv 9:283–291.

Ortiz de Solórzano C, Garcia Rodriguez E, Jones A, Pinkel D, Gray JW, Sudar D, Lockett SJ. 1999. Segmentation of confocal micro-

scope images of cell nuclei in thick tissue sections. J Microsc 193: 212–226.
Ortiz de Solórzano C, Malladi R, Lelievre S, Lockett SJ. 2001. Segmentation of nuclei and cells using membrane related protein markers. J Microsc 201:404–415.
Parvin B, Taylor J, Cong G, O'Keefe MA, Barcellos-Hoff MH. DeepView: A channel for distributed microscopy and informatics. In: Proceedings of the ACM/IEEE SCC99 Conference on High Performance Networking and Computing, Portland, OR, 14–19 November 1999, p 15.
Parvin B, Yang Q, Fontenay G, Henshall R, Barcellos-Hoff MH. 2000. BioSig: an imaging bioinformatic system for studying phenomics. IEEE Computer, 35:65–71.
Petersen OW, Ronnov-Jessen L, Howlett AR, Bissell MJ. 1992. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. Proc Natl Acad Sci USA 89:9064–9068.
Roberts AB, Thompson NL, Heine U, Flanders C, Sporn MB. 1988. Transforming growth factor-beta: possible roles in carcinogenesis. Br J Cancer 57:594–600.
Roskelley CD, Srebrow A, Bissell MJ. 1995. A hierarchy of ECM-mediated signalling regulates tissue-specific gene expression. Curr Op Cell Biol 7:736–747.
Shyamala G, Barcellos-Hoff MH, Toft D, Yang X. 1997. In situ localization of progesterone receptors in normal mouse mammary glands: absence of receptors in the connective and adipose stroma and a heterogeneous distribution in the epithelium. J Steroid Biochem Mol Biol 63:251–259.
Siegal GP, Barsky SH, Terranova VP, Liotta LA. 1981. Stages of neoplastic transformation of human breast tissue as monitored by dissolution of basement membrane components. An immunoperoxidase study. Invas Metast 1:54–70.
Stack MS, Gray RD, Pizzo SV. 1993. Modulation of murine B16F10 melanoma plasminogen activator production by a synthetic peptide derived from the laminin A chain. Cancer Res 53:1998–2004.
Tlsty TD. 1998. Cell-adhesion-dependent influences on genomic instability and carcinogenesis. Curr Opin Cell Biol 10:647–653.
Tokunaga M, Land C, Tokuoka S. 1991. Follow-up studies of breast cancer incidence among atomic bomb survivors. J Rad Res Suppl 32:201–211.
Weaver VM, Petersen OW, Wang F, Larabell CA, Briand P, Damsky C, Bissell MJ. 1997. Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies. J Cell Biol 137:231–245.